

## CLAIMS

1. A system for enhancing an audio reception experience comprising:
  - a visual output device;
  - visual content storage supplying visual content to said visual output device;
  - an audio player operative to play audio content containing non-synthesized voice; and
  - an audio-visual coordinator operative to cause said visual output device to display said visual content in a manner coordinated with said non-synthesized voice.
2. A system according to claim 1 and wherein said audio-visual coordinator is operative to extract phonemes from said voice and to match said phonemes to visemes in said visual content.
3. A system according to claim 1 and wherein said visual content includes at least one image of at least one person speaking.
4. A system according to claim 3 and wherein said at least one image comprises a plurality of images, each representing at least one viseme.
5. A system according to claim 1 and wherein said visual output device comprises a display screen.
6. A system according to claim 1 and wherein said visual output device comprises a three-dimensional animated object.
7. A system according to claim 6 and wherein said three-dimensional animated object is operative to present a plurality of different visemes.
8. A system according to claim 7 and wherein said three-dimensional animated object is operative to present visemes which are time coordinated with

phonemes in said voice.

9. A system according to claim 1 and wherein said visual output device is operative to provide visual cues coordinated with various parameters of said voice.

10. A system according to claim 9 and wherein said various parameters include at least one of intonation, volume, pitch and emphasis.

11. A system for enhancing an audio reception experience comprising:  
a three-dimensional animated visual output device;  
visual content storage supplying visual content to said visual output device;  
an audio player operative to play audio content containing voice; and  
an audio-visual coordinator operative to cause said visual output device to display said visual content in a manner coordinated with said voice.

12. A system according to claim 11 and wherein said audio-visual coordinator is operative to extract phonemes from said voice and to match said phonemes to visemes in said visual content.

13. A system according to claim 11 and wherein said visual content includes at least one image of at least one person speaking.

14. A system according to claim 13 and wherein said at least one image comprises a plurality of images, each representing at least one viseme.

15. A system according to claim 11 and wherein said three-dimensional animated object is operative to present a plurality of different visemes.

16. A system according to claim 15 and wherein said three-dimensional animated object is operative to present visemes which are time coordinated with phonemes in said voice.

17. A system according to claim 11 and wherein said visual output device is operative to provide visual cues coordinated with various parameters of said voice.
18. A system according to claim 17 and wherein said various parameters include at least one of intonation, volume, pitch and emphasis.
19. A system for enhancing an audio reception experience comprising:
  - a visual output device;
  - visual content storage supplying visual content to said visual output device;
  - an audio player operative to play audio content containing voice; and
  - an audio-visual coordinator operative to cause said visual output device to display said visual content in a manner coordinated with said voice, said audio-visual coordinator being operative to extract phonemes from said voice and to match said phonemes to visemes in said visual content.
20. A system according to claim 19 and wherein said visual content includes at least one image of at least one person speaking.
21. A system according to claim 20 and wherein said at least one image comprises a plurality of images, each representing at least one viseme.
22. A system according to claim 19 and wherein said visual output device comprises a display screen.
23. A system according to claim 19 and wherein said visual output device comprises a three-dimensional animated object.
24. A system according to claim 23 and wherein said three-dimensional animated object is operative to present a plurality of different visemes.

25. A system according to claim 24 and wherein said three-dimensional animated object is operative to present visemes which are time coordinated with phonemes in said voice.

26. A system according to claim 19 and wherein said visual output device is operative to provide visual cues coordinated with various parameters of said voice.

27. A system according to claim 26 and wherein said various parameters include at least one of intonation, volume, pitch and emphasis.

28. For use with a visual output device and an audio player operative to play audio content,

an audio reception experience enhancement module comprising:

visual content storage supplying visual content to said visual output device; and

an audio-visual coordinator operative to cause said visual output device to display said visual content in a manner coordinated with said audio content.

29. For use with a three-dimensional animated visual output device and an audio player operative to play audio content,

an audio reception experience enhancement module comprising:

visual content storage supplying visual content to said visual output device; and

an audio-visual coordinator operative to cause said visual output device to display said visual content in a manner coordinated with said audio content.

30. For use with a visual output device and an audio player operative to play audio content,

an audio reception experience enhancement module comprising:

visual content storage supplying visual content to said visual output device; and

an audio-visual coordinator operative to cause said visual output device

to display said visual content in a manner coordinated with said audio content, said audio-visual coordinator being operative to extract phonemes from said audio content and to match said phonemes to visemes in said visual content.

31. Apparatus for generating a visual representation of speech comprising:

a reservoir of viseme profiles storing at least one viseme profile, each viseme profile including a complete set of visemes respectively depicting different speech production positions of a persona, each viseme profile being linked to information identifying its persona;

a phoneme extractor operative to receive a speech input and to derive therefrom a timed sequence of phonemes included therewithin; and

a visual speech representation generator operative to access a viseme profile from said reservoir and to present a visual representation to accompany said speech input, the visual representation including a viseme sequence formed from visemes included in the viseme profile which respectively match the phonemes in said timed sequence, wherein the visual representation generator presents each viseme generally simultaneously with its matching phoneme.

32. Apparatus according to claim 31 and also comprising a user interface operative to prompt a user to define at least one characteristic of at least one telephone communication session and to select at least one viseme profile within said reservoir to be associated with said telephone communicant.

33. Apparatus according to claim 32 and wherein said visual speech representation generator is operative to present a visual representation formed from the viseme profile selected by the user, to accompany a speech input generated in the course of said telephone communication session.

34. Apparatus according to claim 31 wherein said visual speech representation generator comprises apparatus for generating a visual speech representation which is integrally formed with a household appliance.

35. Apparatus according to claim 31 wherein said reservoir of viseme profiles comprises a user interface operative to prompt a user to provide a viseme profile access request including confirmable information identifying a persona whose viseme profile the user wishes to access, and also operative to provide the persona's viseme profile to the user.

36. Apparatus according to claim 35 wherein the user interface and the user communicate via a computer network.

37. Apparatus according to claim 35 wherein said user interface is also operative to impose a charge for providing the persona's viseme profile to the user including obtaining the user's approval therefor before providing the persona's viseme profile to the user.

38. Apparatus according to claim 31 wherein said visual speech representation generator comprises apparatus for generating a visual speech representation which is integrally formed with a goods vending device.

39. Apparatus according to claim 38 wherein said goods vending device comprises a beverage dispensing machine.

40. Apparatus according to claim 31 wherein said visual speech representation generator comprises apparatus for generating a visual speech representation which is integrally formed with a services dispensing device.

41. Apparatus according to claim 40 wherein said services dispensing device comprises an automatic bank teller.

42. Apparatus according to claim 31 wherein said visual speech representation generator is operative to present the visual representation on a display screen of a communication device.

43. Apparatus according to claim 42 wherein the communication device comprises an individual one of the following group of communication devices having display screens: personal digital assistant, cellular telephone such as a third generation cellular telephone, wired telephone, radio, interactive television, beeper device, computer such as a personal computer, portable computer or household computer, television, screenphone, electronic game, and devices having a plurality of physical positions which can be correspond to speech production positions.

44. Apparatus according to claim 31 wherein said reservoir, phoneme extractor and visual speech representation generator are all cached in a telephone.

45. A method for generating a visual representation of speech comprising:  
providing a reservoir of viseme profiles storing at least one viseme profile, each viseme profile including a complete set of visemes respectively depicting different speech production positions of a persona, each viseme profile being linked to information identifying its persona;  
receiving a speech input and deriving therefrom a timed sequence of phonemes included therewithin; and  
accessing a viseme profile from said reservoir and presenting a visual representation to accompany said speech input, the visual representation including a viseme sequence formed from visemes included in the viseme profile which respectively match the phonemes in said timed sequence, wherein each viseme is presented generally simultaneously with its matching phoneme.

46. A method according to claim 45 wherein said step of providing a reservoir comprises, for each of a plurality of personas:

generating a sequence of visual images representing the persona uttering a speech specimen including all visemes in a particular language; and  
identifying from within the sequence of visual images, and storing, a complete set of visemes.

47. A method according to claim 45 wherein said step of providing

comprises storing at least one viseme profile in a first communication device serving a first communicant and, upon initiation of a communication session between the first communicant and a second communicant, transmitting the viseme profile between the first communication device and a second communication device serving the second communicant,

and wherein said step of accessing and presenting comprises presenting, on a screen display associated with the second communication device, a viseme sequence formed from visemes included in the viseme profile transmitted from the first communicant to the second communicant.

48. A method according to claim 47 wherein said step of transmitting comprises sending the viseme profile in near real time via a data channel while a telephone call is in progress.

49. A method according to claim 47 wherein said step of sending employs a multimedia messaging service.

50. Apparatus for generating a visual representation of speech comprising:  
a toy having several speech production positions;  
a speech production position memory associating each phoneme in a language with an individual one of the speech production positions;  
a phoneme extractor operative to receive a speech input, to derive therefrom a timed sequence of phonemes included therewithin, and to derive therefrom, using said speech production position memory, a correspondingly timed sequence of speech production positions respectively corresponding to the phonemes in said timed sequence; and  
a toy speech position controller operative to actuate the toy to adopt said correspondingly timed sequence of speech production positions.

51. A business card comprising:  
a card presenting contact information regarding a bearer of the card including information facilitating access to a viseme profile of the bearer.

52. Stationery apparatus comprising:  
stationery paper including a header presenting contact information for at least one individual including information facilitating access to a viseme profile of at least one individual.
53. A website comprising:  
a web page presenting contact information for at least one individual associated with the website including information facilitating access to a viseme profile of the individual.